

# SOHAIL(NEEL) SARKAR

☎ 437-329-3448

✉ [sohail.sarkar@utoronto.ca](mailto:sohail.sarkar@utoronto.ca)

🌐 [n33levo.github.io](https://n33levo.github.io)

🌐 [linkedin.com/in/n33levo](https://linkedin.com/in/n33levo)

🐙 [github.com/n33levo](https://github.com/n33levo)

## EDUCATION

### University of Toronto

B.Sc. Applied Mathematics and Computer Science

Expected May 2027

Toronto, ON

### Trinity College

B.A. Music, ACTL Classical Guitar

Jun 2021

London, England

## EXPERIENCE

### Sigma Squared

Applied Scientist Intern

Jun 2025 – Present

Boston, MA (remote)

- Engineered an agentic EDA to ETL for success-factors behavior prediction, reduced variance by 27%, & RMSE by 8%.
- Developed Agentic SDK with MCP wrapping, scaling 50+ workflows across data pipelines at S2 and EO Ventures.
- Co-authoring proprietary research on econometric-based performance clustering and calibration using Set and Group theory.

### PSP Investments

Data Science Intern

May – Jul 2025

Montréal, QC

- Architected alpha-signal data processing pipeline answering 3,000+ analyst questions per batch.
- Reinforced post-processing with claimification-based cross-referencing; achieved 100% factual accuracy.
- Built Scala retriever layers for preprocessing & learned query routing; cut batch inference from 20 to 3 min.

### University of Toronto

MLOps Intern

Jan – Apr 2025

Toronto, ON

- Delivered a distributed, hierarchical retriever enabling multi-tenant chat and scalable context routing across teams.
- Deployed RAG framework with CI/CD automaton for indexing, evaluation, and releases; cut release times by 60% (5d→2d).
- Exposed library GPU cluster with REST control-plane for LLM fine-tuning, including job scheduling and telemetry.

### E.J. Pratt Library

Software Engineering Intern

Sept – Dec 2024

Toronto, ON

- Developed version control scheduling; optimized Random Forest to be 6× faster, result & 12% more shift coverage.
- Shipped OAuth-secured ticketing (MongoDB + Redis); reduced IT desk resolution time by 40% (3d→1.8d).

## PROJECTS

🗨️ [TALK TO MY PA](#)

### Fine-Tuned Stock Analysis LLM | Bayesian-LoRA, 4-bit quantization, PyTorch, PEFT, bitsandbytes, CUDA

Nov 2024

- Fine-tuned LLaMA-3.1-8B w 4-bit quantized Bayesian-LoRA on market reports, earnings calls & SEC filings; used Bayesian rank gating to adapt per-layer ranks while reducing compute & memory.

### Recommendation Engine | Neo4j, PyTorch, FastAPI, Dagster, AWS EKS(Kubernetes), Docker

Mar 2024

- Hybrid recommender implicit-feedback ALS on watch history & Neo4j knowledge-graph features with a trained re-ranker; Dockerized FastAPI inference & Dagster refresh pipelines and deployed on EKS w autoscaled endpoints.

### OCR Digitizer Desktop Software | TensorFlow.NET, C#, CRNN/CTC, Beam Search, Emgu CV (OpenCV)

May 2023

- Built OCR software extracting structured data from scans stores in PostgreSQL, exports to Excel; deployed to 30+ supply-chain and law firms across North Bengal as on-premises licensed software, generated CAD \$35K in revenue.

### Modified SIR | immunology, PDEs, SciPy, Pandas, PyMC3, FilterPy

Feb 2022

- Age-stratified SEIRV extension of SIR included Exposed, Vaccinated & heterogeneous risks; solved ODEs in SciPy with Bayesian fit in PyMC3; applied Unscented & Particle Filters for sequential updates of S/E/I/R/V on COVID-19 data.

## TECHNICAL SKILLS

**Languages:** Python, Go, Bash, TypeScript, C/C++, Assembly, SQL

**Frameworks & Libraries:** PyTorch, TensorFlow, CUDA, Hugging Face, LangChain, LlamaIndex, FastAPI, FastMCP, vLLM, Triton, scikit-learn, XGBoost, OpenCV, Streamlit, Tailwind CSS, Agentic SDKs

**Data & Infra:** Spark, Databricks, MLflow, Kafka, Docker, k8s, AWS, REST, GraphQL, Jenkins, Open AI Batch

**Databases:** Neo4j, Pinecone, PostgreSQL, MongoDB, Redis, SQLite